

**Tilburg University**

## **Stochastic games with one big action space in each state**

Tijs, S.H.

*Published in:*  
Methods of Operations Research

*Publication date:*  
1980

[Link to publication in Tilburg University Research Portal](#)

*Citation for published version (APA):*  
Tijs, S. H. (1980). Stochastic games with one big action space in each state. *Methods of Operations Research*, 38, 161-173.

### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



# Stochastic Games with one Big Action Space in Each State

S. H. Tijs, Nijmegen

## Abstract.

Non-cooperative stochastic games with a finite state space, a finite player set, metric action spaces and bounded reward functions are considered under the discounted payoff criterion. It is shown, under certain continuity assumptions with respect to reward functions and transition probability functions, that such a game possesses an  $\epsilon$ -equilibrium point in stationary strategies for all  $\epsilon > 0$ , if at each state at most one of the action spaces is topologically big and all other action spaces are topologically small (finite, compact, precompact metric). In the proofs of the  $\epsilon$ -equilibrium point theorems use is made of appropriate subgames, for which  $\epsilon$ -equilibria exist. The existence of these subgames is established with the aid of a result by Fréchet-Ascoli-Arzela.

## 1. Introduction.

Sequential competitive decision processes, such as stochastic games, play an increasing role as useful models for phenomena in the social sciences. We only mention, here, the papers of Deshmukh & Winston [1], Kirman & Sobel [8] and the book of Friedman [5], in which oligopoly situations are related with stochastic games.

The theory of stochastic games started with the fundamental paper of Shapley [13], in which two-person zero-sum stochastic games with finite state space and finite action spaces were considered. The results of Shapley were extended in various directions; for a survey, see [10].

Non-zero-sum stochastic games were studied in the papers [3], [4], [5], [7], [9], [11], [12], [14], [15], [16], [20] and [22]. In all these papers, in the ( $\epsilon$ -)equilibrium point theorems only topologically small action spaces are considered (finite or compact action spaces, or action spaces with a compact closure). Central in this paper is the problem of the existence of  $\epsilon$ -equilibrium points for all  $\epsilon > 0$  for infinite stage stochastic games, where some of the action spaces may be topologically big, while the state space is finite and where the players want to maximize their total expected rewards over all stages, discounted to the initial stage. The technique used in the proofs of the  $\epsilon$ -equilibrium point theorems, derived in this paper, is the approximation of the game with



suitable subgames with small action spaces. A similar technique was also successful in the papers [17] and [18], where  $\epsilon$ -equilibrium point theorems were derived for non-cooperative games in normal form.

The organization of this paper is as follows. In section 2 the stochastic game model is described and some definitions are given. In section 3  $\delta$ -approximations of stochastic games are introduced and some results are presented, which are used in section 4 for the derivation of the new  $\epsilon$ -equilibrium point theorems.

## 2. Model and notation.

We deal with a dynamic system, which can be in a finite number of states  $1, 2, \dots, z$  and which is governed by  $n$  decision makers (players)  $1, 2, \dots, n$ , who at the decision moments (stages)  $0, 1, 2, \dots$  observe the state of the system and then influence the development of the system by choosing one of their actions available at the observed state. At each stage two things happen:

- (1) each of the players obtains an immediate reward, dependent on the state of the system at stage  $k$  and the chosen actions
- (2) the system jumps to one of the states with probabilities also dependent on the state at stage  $k$  and the chosen actions, which new state is observed at the next stage  $k+1$ .

In the following let us denote

- the state space  $\{1, 2, \dots, z\}$  by  $S$ ,
- the set of players  $\{1, 2, \dots, n\}$  by  $I$ ,
- the set of stages  $\{0, 1, 2, \dots\}$  by  $K$ ,
- the non-empty set of pure actions for player  $i \in I$  in state  $s \in S$  by  $A_{is}$ ,
- the immediate reward for player  $i$  in state  $s$ , if the players choose pure actions  $a_1 \in A_{1s}, a_2 \in A_{2s}, \dots, a_n \in A_{ns}$  by  $r_{is}(a_1, a_2, \dots, a_n)$  or  $r_{is}(a)$  and
- the probability that the system is in state  $t \in S$  at stage  $k+1$  by  $p_{st}(a_1, \dots, a_n)$ , if the system is in state  $s \in S$  at stage  $k \in K$  and the players choose actions  $a_1, \dots, a_n$ .

In the following  $\beta$  is a fixed real number in  $(0, 1)$ . We will assume that a reward of  $r$  units, received in stage  $k \in K$ , for a player has worth  $\beta^k r$  at stage 0, and we call this the discounted reward. We assume that the players want to maximize their total discounted reward.

Summarizing and specifying what preceeded, we obtain the following



DEFINITION 1. A stochastic game  $\Gamma$  is a five-tuple

$\langle S, \{A_{is} : i \in I, s \in S\}, \{r_{is} : i \in I, s \in S\}, \{p_{st} : s, t \in S\}, \beta \rangle$ , where

$I$  and  $S$  are as above

$\equiv A_{is}$  is a non-empty metric space (the pure action space for player  $i$  in state  $s$ ).

$\equiv r_{is}$  is a bounded real-valued Borel-measurable function on  $\prod_{i \in I} A_{is}$  (the reward function for player  $i$  in state  $s$ )

$\equiv p_{st}$  is a Borel-measurable function on  $\prod_{i \in I} A_{is}$  for each  $s, t \in S$  (transition probability function) such that

$$p_{st}(a) \geq 0 \text{ and } \sum_{t \in S} p_{st}(a) = 1 \text{ for all } a \in \prod_{i \in I} A_{is}.$$

$\equiv \beta \in (0, 1)$  (the discount factor).

In view of the measurability assumptions in the above definition, we can also take mixed actions into consideration. Here a mixed action for player  $i$  at state  $s$  in the game  $\Gamma$  is a probability measure on the family of Borel sets  $A_{is}$  of  $A_{is}$ . The set of all such mixed actions is denoted by  $\Pi_{is}$ . Using such a mixed action  $\pi_{is} \in \Pi_{is}$  implies that, for each  $U \in A_{is}$ , the probability that a pure action, lying in  $U$ , is chosen, equals  $\pi_{is}(U)$ . The elements of  $\Pi_i = \prod_{s \in S} \Pi_{is}$  are called stationary strategies for player  $i$ .

Using a stationary strategy  $\pi = (\pi_1, \pi_2, \dots, \pi_n)$  means that in all stages player  $i$  uses his mixed action  $\pi_s$ , if the system is in state  $s$ .

*In this paper we only consider stationary strategies* (see section 4, remark 1). The set  $\Pi = \prod_{i \in I} \Pi_i$  is called the outcome space of the game  $\Gamma$ .

In a natural way the reward functions  $r_{is}$  and transition probability functions  $p_{st}$  determine functions  $R_{is}$  and  $P_{st}$  on the outcome space as follows. For each outcome  $\pi = (\pi_1, \pi_2, \dots, \pi_n) \in \Pi$ ,  $s, t \in S$  and  $i \in I$  we define

$$R_{is}(\pi) = \int r_{is}(a_1, a_2, \dots, a_n) d\pi_{1s}(a_1) d\pi_{2s}(a_2) \dots d\pi_{ns}(a_n)$$

$$P_{st}(\pi) = \int p_{st}(a_1, a_2, \dots, a_n) d\pi_{1s}(a_1) d\pi_{2s}(a_2) \dots d\pi_{ns}(a_n).$$

$R_{is}(\pi)$  and  $P_{st}(\pi)$ , respectively, are the expected immediate reward for player  $i$  in state  $s$  and the expected probability that the system is in the next stage in state  $t$ , given that the system is presently in state  $s$ , and given that the players together choose the outcome  $\pi$ .

The reward vector  $(R_{i1}(\pi), \dots, R_{iz}(\pi))$  for player  $i$ , corresponding to



outcome  $\pi$ , will be denoted by  $R_i(\pi)$ . Note that  $P(\pi) = [P_{st}(\pi)]_{s=1, t=1}^z$  is a stochastic  $z \times z$ -matrix, describing the transition probabilities of the system, given the outcome  $\pi$ .

If in the initial stage (stage 0) the system is in  $s$  and if the players choose outcome  $\pi$ , then the  $\beta$ -discounted expected immediate reward at stage  $k$  for player  $i$  equals the  $s$ -th coordinate of the vector  $\beta^k P(\pi)^k R_i(\pi)$  in  $\mathbb{R}^z$ . The total  $\beta$ -discounted expected payoff for player  $i$ , if the system starts in  $s$  and if the players choose  $\pi$ , is given by the  $s$ -th coordinate of the vector  $\sum_{k=0}^{\infty} \beta^k P(\pi)^k R_i(\pi)$  in  $\mathbb{R}^z$ . In the sequel we use the following notation:  $v_i(\pi) = \sum_{k=0}^{\infty} \beta^k P(\pi)^k R_i(\pi)$ ,  $v_i(\pi) = (v_{i1}(\pi), v_{i2}(\pi), \dots, v_{iz}(\pi))$ . For  $a = (a_1, \dots, a_{j-1}, a_j, a_{j+1}, \dots, a_n) \in \prod_{i \in I} A_i$  and  $b \in A_j$  the vector  $(a_1, \dots, a_{j-1}, a_{j+1}, \dots, a_n)$  is denoted by  $a_{-j}$  and the vector  $(a_1, \dots, a_{j-1}, b, a_{j+1}, \dots, a_n)$  by  $(a_{-j} \uparrow b)$ . Similarly, for  $\pi = (\pi_1, \pi_2, \dots, \pi_n) \in \Pi$  and  $\rho \in \Pi_j$  the outcome  $(\pi_1, \pi_2, \dots, \pi_{j-1}, \rho, \pi_{j+1}, \dots, \pi_n)$  is denoted by  $(\pi_{-j} \uparrow \rho)$ .

DEFINITION 2. Let  $\varepsilon \geq 0$  and let  $\hat{\pi} \in \Pi$ . We say that  $\hat{\pi}$  is an  $\varepsilon$ -equilibrium point for the stochastic game  $\Gamma$ , if for each  $i \in I$  and each  $\pi_i \in \Pi_i$  we have

$$(2.1) \quad v_i(\hat{\pi}_{-i} \uparrow \pi_i) \leq v_i(\hat{\pi}) + \varepsilon 1_z.$$

(Here  $1_z$  is the vector in  $\mathbb{R}^z$  with all coordinates equal to 1). 0-equilibrium points are also called equilibrium points.

The main problem in this paper is that of the existence of  $\varepsilon$ -equilibrium points for all  $\varepsilon > 0$ . For zero-sum two-person stochastic games this problem coincides with the problem of the existence of a value (cf. [17] p. 756).

### 3. $\delta$ -Approximations.

In the sequel, use is made of the norms on the space of  $z \times z$ -matrices of real numbers and on  $\mathbb{R}^z$ , defined by

$$||A|| = \max_{t=1}^z \left\{ \sum_{s=1}^z |a_{st}| : s = 1, 2, \dots, z \right\}, \text{ where } A \text{ is the } z \times z\text{-matrix}$$

$$[a_{st}]_{s=1, t=1}^z$$

$$||(x_1, x_2, \dots, x_z)|| = \max_s |x_s| \text{ for } x = (x_1, \dots, x_z) \in \mathbb{R}^z.$$



These norms satisfy the pleasant product properties:

$$(3.1) \quad ||AB|| \leq ||A|| ||B||, \quad ||Ax|| \leq ||A|| ||x|| \text{ for all } z \times z\text{-matrices } A \text{ and } B \text{ and } x \in \mathbb{R}^z.$$

For a stochastic  $z \times z$ -matrix  $P = [p_{st}]_{s=1, t=1}^z$ , for which  $p_{st} \geq 0$  for all  $s, t \in \{1, 2, \dots, z\}$  and  $\sum_{t=1}^z p_{st} = 1$  for all  $s \in \{1, 2, \dots, z\}$ , the following properties hold:

$$(3.2) \quad ||P^k|| = 1 \text{ for all } k \in \mathbb{N}$$

$$(3.3) \quad \sum_{k=0}^{\infty} \beta^k P^k \text{ equals } (I - \beta P)^{-1} \text{ if } \beta \in (0, 1) \text{ and}$$

$$(3.4) \quad || \sum_{k=0}^{\infty} \beta^k P^k || = (1 - \beta)^{-1}.$$

The following lemma is useful, if for two different outcomes one wants to compare the corresponding total  $\beta$ -discounted expected payoff vectors.

Lemma 1. Let  $P$  and  $Q$  be stochastic  $z \times z$ -matrices,  $x, y \in \mathbb{R}^z$  and  $\beta \in (0, 1)$ .

Then

$$(3.5) \quad || \sum_{k=0}^{\infty} \beta^k P^k x - \sum_{k=0}^{\infty} \beta^k Q^k y || \leq (1 - \beta)^{-1} ||x - y|| + \beta(1 - \beta)^{-2} ||P - Q|| ||x||.$$

Proof. First we note that

$$(3.6) \quad \sum_{k=0}^{\infty} \beta^k P^k x - \sum_{k=0}^{\infty} \beta^k Q^k y = \sum_{k=0}^{\infty} \beta^k Q^k (x - y) + \left( \sum_{k=0}^{\infty} \beta^k P^k - \sum_{k=0}^{\infty} \beta^k Q^k \right) x.$$

For the first term on the right-hand side of (3.6), we have, in view of (3.1) and (3.4),

$$(3.7) \quad || \sum_{k=0}^{\infty} \beta^k Q^k (x - y) || \leq || \sum_{k=0}^{\infty} \beta^k Q^k || ||x - y|| = (1 - \beta)^{-1} ||x - y||$$

and for the second term on the right-hand side of (3.6), we obtain, using (3.3), (3.1) and (3.4):

$$\begin{aligned} (3.8) \quad & || \left( \sum_{k=0}^{\infty} \beta^k P^k - \sum_{k=0}^{\infty} \beta^k Q^k \right) x || = || ((I - \beta P)^{-1} - (I - \beta Q)^{-1}) x || = \\ & = || (I - \beta P)^{-1} \beta (P - Q) (I - \beta Q)^{-1} x || \leq \\ & \leq \beta || (I - \beta P)^{-1} || || (I - \beta Q)^{-1} || ||P - Q|| ||x|| = \beta(1 - \beta)^{-2} ||P - Q|| ||x||. \end{aligned}$$

Combining (3.6), (3.7) and (3.8) yields (3.5) □

Let  $\Gamma$  be the stochastic game

$$\langle S, \{A_{is} : i \in I, s \in S\}, \{r_{is} : i \in I, s \in S\}, \{p_{st} : s, t \in S\}, \beta \rangle$$



and put  $\rho(\Gamma) = \sup\{|r_{is}(a)| : i \in I, s \in S, a \in \prod_{i \in I} A_{is}\} \in [0, \infty)$ . Let  $\Gamma'$  be the subgame

$$\langle S, \{B_{is} : i \in I, s \in S\}, \{r'_{is} : i \in I, s \in S\}, \{p'_{st} : s, t \in S\}, \beta \rangle,$$

where  $B_{is}$  is a Borel subset of  $A_{is}$  and where  $r'_{is}$  and  $p'_{st}$  are the restrictions of the maps  $r_{is}$  and  $p_{st}$ , respectively, to the subset  $\prod_{i \in I} B_{is}$  of  $\prod_{i \in I} A_{is}$ .

DEFINITION 3. Let  $\delta \geq 0$ . We will say that the subgame  $\Gamma'$  of  $\Gamma$  is a  $\delta$ -approximation of the game  $\Gamma$  if there exists for each  $(i, s) \in I \times S$  a measurable map  $\phi_{is} : A_{is} \rightarrow B_{is}$  such that for all  $b \in \prod_{i \in I} B_{is}$ ,  $a_i \in A_{is}$  and  $t \in S$  we have

$$(3.9) \quad |r'_{is}(b_{-i} \uparrow \phi_{is}(a_i)) - r_{is}(b_{-i} \uparrow a_i)| \leq \frac{1}{2}(1-\beta)\delta,$$

$$(3.10) \quad |p'_{st}(b_{-i} \uparrow \phi_{is}(a_i)) - p_{st}(b_{-i} \uparrow a_i)| \leq \frac{1}{2}\beta^{-1}(1-\beta)^2 z^{-1}(1+\rho(\Gamma))^{-1}\delta.$$

Let us denote the set of mixed actions of player  $i$  in state  $s$  in the game  $\Gamma'$  by  $\Pi'_{is}$ , the set of stationary strategies of player  $i$  in  $\Gamma'$  by  $\Pi'_i$  and the set of outcomes by  $\Pi'$ . Note that  $\Pi'$  can be identified with a subset of  $\Pi$  in a natural way. The maps  $\phi_{is} : A_{is} \rightarrow B_{is}$  induce maps  $\psi_{is} : \Pi_{is} \rightarrow \Pi'_{is}$  and  $\psi_i : \Pi_i \rightarrow \Pi'_i$  as follows:

$$\begin{aligned} \psi_{is}(\pi_{is})(U) &= \pi_{is}(\phi_{is}^{-1}(U)) \text{ for each } \pi_{is} \in \Pi_{is} \text{ and each} \\ &\quad \text{Borel set } U \text{ in } B_{is} \\ \psi_i(\pi_i) &= (\psi_{i1}(\pi_{i1}), \dots, \psi_{iz}(\pi_{iz})) \text{ for } (\pi_{i1}, \dots, \pi_{iz}) \in \Pi_i. \end{aligned}$$

LEMMA 2. Let  $\delta \geq 0$ . Let  $\Gamma'$  be a  $\delta$ -approximation of  $\Gamma$  via the family of maps  $\{\phi_{is} : i \in I, s \in S\}$ . Let  $\psi_i$ ,  $i \in I$ , be as above. Then for all  $i \in I$ , all  $\pi' \in \Pi'$  and all  $\pi_i \in \Pi_i$  we have

$$||v'_i(\pi'_{-i} \uparrow \psi_i(\pi_i)) - v_i(\pi'_{-i} \uparrow \pi_i)|| \leq \delta$$

(Here  $v'_i(\pi')$  is the total  $\beta$ -discounted expected payoff vector, corresponding to  $\pi'$ , in  $\Gamma'$ ).

PROOF. Take  $i \in I$ ,  $\pi' \in \Pi'$ ,  $\pi_i = (\pi_{i1}, \dots, \pi_{is}, \dots, \pi_{iz}) \in \Pi_i$  and  $s \in S$ . It follows from theorem C on p.163 in HALMOS [6], that

$$\begin{aligned} \int_{B_{is}} r'_{is}(b_1, \dots, b_n) d\psi_{is}(\pi_{is})(b_i) &= \int_{A_{is}} r_{is}(b_1, \dots, b_{i-1}, \phi_{is}(a_i), \\ &\quad b_{i+1}, \dots, b_n) d\pi_{is}(a_i). \end{aligned}$$

Hence, by (3.9), we have



$$\begin{aligned}
& \left| \int_{B_{is}} r'_{is}(b) d\psi_{is}(\pi_{is})(b_i) - \int_{A_{is}} r_{is}(b_{-i} \uparrow a_i) d\pi_{is}(a_i) \right| = \\
& \left| \int_{A_{is}} r_{is}(b_{-i} \uparrow \phi_{is}(a_i)) - r_{is}(b_{-i} \uparrow a_i) d\pi_{is}(a_i) \right| \leq \frac{1}{2}(1-\beta)\delta.
\end{aligned}$$

But then

$$(3.11) \quad |R'_{is}(\pi'_{-is} \uparrow \psi_{is}(\pi_{is})) - R_{is}(\pi'_{-is} \uparrow \pi_{is})| \leq \frac{1}{2}(1-\beta)\delta$$

where  $\pi'_{-is} = (\pi'_{1s}, \dots, \pi'_{i-1,s}, \pi'_{i+1,s}, \dots, \pi'_{ns})$ .

In a similar way, it follows from (3.10) that for all  $s, t \in S$

$$(3.12) \quad |p'_{st}(\pi'_{-is} \uparrow \psi_{is}(\pi_{is})) - p_{st}(\pi'_{-is} \uparrow \pi_{is})| \leq \frac{1}{2}\beta^{-1}(1-\beta)^2 z^{-1}(1+\rho(\Gamma))^{-1}\delta.$$

Now we apply lemma 1, where the roles of  $P$ ,  $Q$ ,  $x$  and  $y$  are played by  $[p_{st}(\pi'_{-is} \uparrow \pi_{is})]_{s=1, t=1}^z$ ,  $[p'_{st}(\pi'_{-is} \uparrow \psi_{is}(\pi_{is}))]_{s=1, t=1}^z$ ,  $R_i(\pi'_{-i} \uparrow \pi_i)$  and  $R'_i(\pi'_{-i} \uparrow \psi_i(\pi_i))$ , respectively.

Using (3.11), (3.12) and the inequality  $||R_i(\pi'_{-i} \uparrow \pi_i)|| \leq \rho(\Gamma)$ , we obtain

$$\begin{aligned}
& ||v'_i(\pi'_{-i} \uparrow \psi_i(\pi_i)) - v_i(\pi'_{-i} \uparrow \pi_i)|| = \\
& = ||\sum \beta^k (P'(\pi'_{-i} \uparrow \psi_i(\pi_i)))^k R'_i(\pi'_{-i} \uparrow \psi_i(\pi_i)) + \\
& - \sum \beta^k (P(\pi'_{-i} \uparrow \pi_i))^k R_i(\pi'_{-i} \uparrow \pi_i)|| \leq \frac{1}{2}\delta + \frac{1}{2}\delta(1+\rho(\Gamma))^{-1}\rho(\Gamma) \leq \delta. \quad \square
\end{aligned}$$

The importance of  $\delta$ -approximations lies in the following

**THEOREM 1.** Let  $\delta \geq 0$ ,  $\epsilon \geq 0$ . Let  $\Gamma'$  be a  $\delta$ -approximation of the game  $\Gamma$  (via  $\{\phi_{is} : i \in I, s \in S\}$ ). Let  $\pi' \in \Pi'$  be an  $\epsilon$ -equilibrium point of the game  $\Gamma'$ . Then  $\pi'$  is an  $(\epsilon+\delta)$ -equilibrium point of  $\Gamma$ .

**PROOF.** Take  $i \in I$  and  $\pi_i \in \Pi_i$ . We have to prove that

$$(3.13) \quad v_i(\pi'_{-i} \uparrow \pi_i) \leq v_i(\pi') + (\epsilon+\delta)1_z$$

holds. In view of lemma 2, we have

$$(3.14) \quad v_i(\pi'_{-i} \uparrow \pi_i) \leq v'_i(\pi'_{-i} \uparrow \psi_i(\pi_i)) + \delta 1_z.$$

Since  $\pi'$  is an  $\epsilon$ -equilibrium point in  $\Gamma'$ , we also have

$$(3.15) \quad v'_i(\pi'_{-i} \uparrow \psi_i(\pi_i)) \leq v'_i(\pi') + \epsilon 1_z.$$

Combining (3.14) and (3.15) yields (3.13).  $\square$

Theorem 1 will be very useful in proving the existence of  $\epsilon$ -equilibrium points for various classes of games. If, namely, for a stochastic



game  $\Gamma$  for each  $\delta > 0$  a  $\delta$ -approximation can be found, which has  $\varepsilon$ -equilibria for each  $\varepsilon > 0$ , then the game  $\Gamma$  also has  $\varepsilon$ -equilibrium points for each  $\varepsilon > 0$ . The following lemma, which is a well-known extension of Fréchet of the Arzela-Ascoli theorem (cf. [2] p. 266, theorem 4.7 and p. 382), will be useful in finding suitable  $\delta$ -approximations.

**LEMMA 3.** Let  $E$  be a compact topological space, let  $F$  be a bounded and equicontinuous family of maps from  $E$  into  $\mathbb{R}^m$  and let  $\eta > 0$ . Then there exists a finite subfamily  $G$  of  $F$  such that  $G$   $\eta$ -covers  $F$  i.e.

$$(3.16) \quad \forall_{f \in F} \exists_{g \in G} \forall_{e \in E} [||f(e) - g(e)|| \leq \eta].$$

The finite version of lemma 3 is given in the following

**LEMMA 4.** Let  $E$  be a finite set, let  $F$  be a bounded family of maps from  $E$  into  $\mathbb{R}^m$  and let  $\eta > 0$ . Then there exists a finite subfamily  $G$  of  $F$  such that (3.16) holds.

#### 4. Some new $\varepsilon$ -equilibrium point theorems for stochastic games.

We start with a

**DEFINITION 4.** Let  $\Gamma$  be an  $n$ -person stochastic game and  $k \in \{1, 2, \dots, n\}$ . Then we call  $\Gamma$  a  $k$ -finite ( $k$ -compact) game if for each state  $s$  at least  $k$  of the  $n$  action spaces  $\{A_{is} : i \in I\}$  are finite (compact).

An  $n$ -finite game is also called a finite game and an  $n$ -compact game is called a compact game.

For later use, we recall now two well-known equilibrium point theorems.

**THEOREM 2.** (cf. [4], [12] p.13, [14] p.1932 and [9]).

Each finite stochastic game  $\Gamma$  possesses an equilibrium point.

**THEOREM 3.** (cf. [16], [15] p.601, [3] p.146, [22] p.31).

Let  $\Gamma$  be a compact stochastic game and suppose that the reward functions  $r_{is}$  and the transition probability functions  $p_{st}$  are continuous for all  $i \in I$  and  $s, t \in S$ . Then  $\Gamma$  possesses an equilibrium point.

Now we are ready to formulate some new  $\varepsilon$ -equilibrium point theorems.

**THEOREM 4.** Let  $\Gamma$  be an  $(n-1)$ -finite stochastic game. Then  $\Gamma$  possesses an  $\varepsilon$ -equilibrium point for each  $\varepsilon > 0$ .

**PROOF.** Let  $\varepsilon > 0$ . Take  $\eta > 0$  such that

$$(4.1) \quad \max\{2\eta(1-\beta)^{-1}, 2\eta\beta(1-\beta)^{-2}z(1+\rho(\Gamma))\} \leq \varepsilon.$$

(a) We start with the construction of a finite subgame  $\Gamma'$  of  $\Gamma$  with ac-



tion sets  $B_{is}$  ( $i \in I, s \in S$ ) and measurable maps  $\phi_{is} : A_{is} \rightarrow B_{is}$  such that  $\Gamma'$  is an  $\varepsilon$ -approximation of  $\Gamma$  via the maps  $\{\phi_{is} : i \in I, s \in S\}$ . For all  $(i,s) \in I \times S$  such that  $A_{is}$  is a finite set, we take  $B_{is} = A_{is}$  and  $\phi_{is} : A_{is} \rightarrow B_{is}$  the identity map. For infinite action sets we proceed as follows. Let  $A_{is}$  be an infinite set. Then  $A_{-is} = \prod_{j \neq i} A_{js}$  is a finite set, because  $\Gamma$  is an  $(n-1)$ -finite game. We consider the bounded family

$$F_{is} = \{a_{-is} \mapsto (r_{is}(a_{-is} \uparrow x), p_{s1}(a_{-is} \uparrow x), \dots, p_{sz}(a_{-is} \uparrow x)) : x \in A_{is}\}$$

of maps from the finite set  $A_{-is}$  into  $\mathbb{R}^{z+1}$ , where each function corresponds to an element of  $A_{is}$ . Now lemma 4 implies that we can find a finite subset  $B_{is}$  of  $A_{is}$  such that the subfamily

$$G_{is} = \{a_{-is} \mapsto (r_{is}(a_{-is} \uparrow y), p_{s1}(a_{-is} \uparrow y), \dots, p_{sz}(a_{-is} \uparrow y)) : y \in B_{is}\}$$

of  $F_{is}$   $\eta$ -covers  $F_{is}$ . Let the elements of the finite set  $B_{is}$  be  $y_1, y_2, \dots, y_p$ . We define  $\phi_{is} : A_{is} \rightarrow B_{is}$  as follows. For each  $x \in A_{is}$  let  $\phi_{is}(x) = y_r$  if  $r$  is the smallest number in  $\{1, 2, \dots, p\}$  for which the following inequalities hold:

$$(4.2) \quad |r_{is}(a_{-is} \uparrow y_r) - r_{is}(a_{-is} \uparrow x)| \leq \eta$$

$$(4.3) \quad |p_{st}(a_{-is} \uparrow y_r) - p_{st}(a_{-is} \uparrow x)| \leq \eta$$

for all  $a_{-is} \in A_{-is}$  and all  $t \in S$ .

Such an  $r$  exists because  $G_{is}$   $\eta$ -covers  $F_{is}$ .

Obviously,  $\phi_{is}$  is measurable. It follows from (4.1), (4.2) and (4.3) that for all  $a_{-is} \in A_{-is} = B_{-is}$  and all  $t \in S$ :

$$(4.5) \quad |r_{is}(a_{-is} \uparrow \phi_{is}(x)) - r_{is}(a_{-is} \uparrow x)| \leq \eta \leq \frac{1}{2}(1-\beta)\varepsilon$$

$$(4.6) \quad |p_{st}(a_{-is} \uparrow \phi_{is}(x)) - p_{st}(a_{-is} \uparrow x)| \leq \eta \leq \frac{1}{2}\beta(1-\beta)^2 z^{-1} (1+\rho(\Gamma))^{-1} \varepsilon.$$

Comparing (4.5) and (4.6) with (3.9) and (3.10), we see that the game  $\Gamma'$  with action sets  $B_{is}$  as chosen above is the desired  $\varepsilon$ -approximation of  $\Gamma$  via  $\{\phi_{is} : i \in I, s \in S\}$ .

(b) By theorem 2 the finite game  $\Gamma'$  has an equilibrium point. Then  $\Gamma$  has an  $\varepsilon$ -equilibrium point, in view of theorem 1.  $\square$

The following example shows that there exist  $(n-2)$ -finite stochastic games, for which there exists no  $\varepsilon$ -equilibrium point for all sufficiently small positive  $\varepsilon$ .



EXAMPLE 1. Let  $\Gamma$  be the 2-person zero-sum stochastic game  $\langle S, \{A_{11}, A_{21}\}, \{r_{11}, r_{21}\}, \{p_{11}\}, \beta \rangle$  with  $S = \{1\}$  (one state),  $A_{11} = A_{21} = \mathbb{N}$ ,  $\beta \in (0, 1)$ ,  $p_{11}(a_1, a_2) = 1$  for all  $(a_1, a_2) \in A_{11} \times A_{21}$  and

$$r_{11}(a_1, a_2) = -r_{21}(a_1, a_2) = \begin{cases} 1 & \text{if } a_1 \geq a_2 \\ 0 & \text{if } a_1 < a_2 \end{cases}.$$

(Cf. WALD [21] p.282).

Then the lower value  $\sup_{\pi_1 \in \Pi_1} \inf_{\pi_2 \in \Pi_2} v(\pi_1, \pi_2)$  equals 0 and the upper value

$\inf_{\pi_2 \in \Pi_2} \sup_{\pi_1 \in \Pi_1} v(\pi_1, \pi_2)$  equals  $(1-\beta)^{-1}$ . This implies that  $\Gamma$  has no  $\epsilon$ -equili-

brium points if  $\epsilon < \frac{1}{2}(1-\beta)^{-1}$ .

Now we want to look at  $(n-1)$ -compact stochastic games, for which the following continuity conditions hold:

(C) for all  $(i, s) \in I \times S$  for which  $A_{is}$  is compact, the family  $\{a_{is} \mapsto (r_i, p_{s1}, p_{s2}, \dots, p_{sz})(x \uparrow a_{is}) : x \in A_{-is}\}$  of maps from  $A_{is}$  into  $\mathbb{R}^{z+1}$  is an equicontinuous family.

THEOREM 5. Let  $\Gamma$  be an  $(n-1)$ -compact game, satisfying property (C). Then  $\Gamma$  possesses an  $\epsilon$ -equilibrium point for each  $\epsilon > 0$ .

PROOF. Let  $\epsilon > 0$ . Take  $\eta > 0$  such that

$$(4.7) \quad \max\{4\eta(1-\beta)^{-1}, 4\eta\beta(1-\beta)^{-2}z(1+\rho(\Gamma))\} \leq \epsilon.$$

We will construct, for each  $(i, s) \in I \times S$ , a subset  $B_{is}$  of  $A_{is}$  and a measurable function  $\phi_{is} : A_{is} \rightarrow B_{is}$  such that the subgame  $\Gamma$  with the action spaces  $\{B_{is} : i \in I, s \in S\}$  is an  $(n-1)$ -finite subgame, which  $\frac{1}{2}\epsilon$ -approximates  $\Gamma$  via the maps  $\{\phi_{is} : i \in I, s \in S\}$ .

For all  $(i, s) \in I \times S$ , with  $A_{is}$  non-compact, we take  $A_{is} = B_{is}$  and  $\phi_{is} : A_{is} \rightarrow B_{is}$  the identity map.

Now let  $(i, s) \in I \times S$  be such that  $A_{is}$  is compact. Denote the function  $a_{is} \mapsto (r_i, p_{s1}, p_{s2}, \dots, p_{sz})(x \uparrow a_{is})$  on  $A_{is}$ , where  $x \in A_{-is}$ , by  $f_x$ ; and the ball of all points in  $A_{is}$  with distance to  $a \in A_{is}$  at most  $\delta$  by  $B(a, \delta)$ . The equicontinuity property of the family  $\{f_x : x \in A_{-is}\}$  and the compactness of  $A_{is}$  imply that there is a finite subset  $B_{is}$  of  $A_{is}$  such that

$$(4.8) \quad \forall b \in B_{is} \quad \exists \delta(b) > 0 \quad \forall a_{is} \in B(b, \delta(b)) \quad \forall x \in A_{-is} \quad [|f_x(a_{is}) - f_x(b)| \leq \eta]$$

and

$$\bigcup_{b \in B_{is}} B(b, \delta(b)) = A_{is}.$$



To construct the measurable map  $\phi_{is} : A_{is} \rightarrow B_{is}$ , suppose that  $B_{is} = \{p_1, p_2, \dots, p_r\}$ . For  $a_{is} \in A_{is}$  put  $\phi_{is}(a_{is}) = p_k$  if  $a_{is} \in B(p_k, \delta(b_k))$  and  $a_{is} \notin \bigcup_{m=1}^{k-1} B(p_m, \delta(b_m))$ . Then  $\phi_{is}$  satisfies the desired properties. Now it

follows from (4.7) and (4.8) that for all  $b_{-i} \in B_{-is}$ ,  $a_{is} \in A_{is}$  and  $t \in S$  we have

$$(4.9) \quad |r_{is}(b_{-i} \uparrow \phi_{is}(a_{is})) - r_{is}(b_{-i} \uparrow a_{is})| \leq \eta \leq \frac{1}{4}\epsilon(1-\beta),$$

$$(4.10) \quad |p_{st}(b_{-i} \uparrow \phi_{is}(a_{is})) - p_{st}(b_{-i} \uparrow a_{is})| \leq \eta \leq \frac{1}{4}\epsilon\beta^{-1}(1-\beta)^2 z^{-1}(1+\rho(\Gamma))^{-1}.$$

By (4.9) and (4.10) the constructed game  $\tilde{\Gamma}$  is a  $\frac{1}{4}\epsilon$ -approximation of  $\Gamma$ . Since  $\tilde{\Gamma}$  is an  $(n-1)$ -finite game,  $\tilde{\Gamma}$  possesses an  $\frac{1}{4}\epsilon$ -equilibrium point by theorem 4; this point is an  $\epsilon$ -equilibrium point of  $\Gamma$ , in view of theorem 1.  $\square$

Now we look at  $(n-1)$ -compact games  $\Gamma$  satisfying the following continuity conditions:

(D.1) All reward functions  $r_{is}$  and transition probability functions  $p_{st}$  ( $i \in I, s \in S, t \in S$ ) are continuous.

(D.2) For each  $(i, s) \in I \times S$  with  $A_{is}$  non-compact the family

$$F_{is} = \{a_{-is} \mapsto (r_i, p_{s1}, \dots, p_{sz})(a_{-is} \uparrow y) : y \in A_{is}\}$$

is an equicontinuous family of maps from  $A_{-is}$  into  $\mathbb{R}^{z+1}$ .

**THEOREM 6.** Let  $\Gamma$  be an  $(n-1)$ -compact game, satisfying (D.1) and (D.2).

Then  $\Gamma$  possesses an  $\epsilon$ -equilibrium point for each  $\epsilon > 0$ .

**PROOF.** Let  $\epsilon > 0$ . Take  $\eta > 0$  satisfying (4.1). In view of theorems 3 and 1, it is sufficient to show that  $\Gamma$  possesses an  $n$ -compact  $\epsilon$ -approximation  $\Gamma'$  with continuous reward functions and transition probability functions.

We construct the action spaces  $\{B_{is} : i \in I, s \in S\}$  for  $\Gamma'$  as follows. If  $A_{is}$  is compact, take  $B_{is} = A_{is}$  and  $\phi_{is}$  the identity map. For  $A_{is}$  non-compact, condition (D.2) and lemma 3 imply the existence of a finite subset  $B_{is}$  of  $A_{is}$  such that the family

$$G_{is} = \{a_{-is} \mapsto (r_i, p_{s1}, \dots, p_{sz})(a_{-is} \uparrow y) : y \in B_{is}\}$$

$\eta$ -covers the family  $F_{is}$ . As in the proof of theorem 4, we can, in a similar way, also find a measurable  $\phi_{is} : A_{is} \rightarrow B_{is}$ , such that (4.5) and (4.6) hold. Now  $\Gamma'$ , with the action spaces  $\{B_{is} : i \in I, s \in S\}$  is a compact game,  $\epsilon$ -approximating  $\Gamma$  via  $\{\phi_{is} : i \in I, s \in S\}$ .  $\square$



We conclude this paper with some  
REMARKS.

1. In the foregoing we restricted our attention to stationary strategies. This is motivated by the fact that, as in Vrieze [20], it can be shown that an  $\epsilon$ -equilibrium point in the class of stationary strategies is also an  $\epsilon$ -equilibrium point in the class of history dependent strategies (cf. [5] p.281).
2. For the stochastic games, described in theorems 4, 5 and 6, there exists for each  $\epsilon > 0$  an  $\epsilon$ -equilibrium point  $(\pi_1, \pi_2, \dots, \pi_n)$  such that, for each  $i \in I$  and  $s \in S$ , the mixed action  $\pi_{is}$  uses only a finite number of pure actions. This follows from the proofs of these theorems.
3. As in theorem 5, one can prove that an  $(n-1)$ -precompact stochastic game, for which the families of maps, described in (C), are uniform equicontinuous families, possesses an  $\epsilon$ -equilibrium point for each  $\epsilon > 0$ .
4. Also Whitt used in [22] an approximation technique to obtain  $\epsilon$ -equilibrium point theorems. But in his theorems, the action spaces all have compact closures, so they are topologically small.
5. In a forthcoming paper [19], O.J. Vrieze and the author extend some of the results of this paper in various directions.

Acknowledgement.

The author wishes to thank O.J. Vrieze for helpful discussions.

REFERENCES.

- [1] DESHMUKH, S.D. & WINSTON, W. A zero-sum stochastic game model of duopoly, Int.J. of Game Theory 7, 19-30 (1978).
- [2] DUNFORD, N. & SCHWARTZ, J.T. Linear Operators I, Interscience, New York (1958).
- [3] FEDERGRUEN, A. Markovian control problems, Math. Centrum, Amsterdam (1978).
- [4] FINK, A.M. Equilibrium in a stochastic n-person game, J. Sci. Hiroshima Univ. Ser. A-I 28, 89-93 (1964).
- [5] FRIEDMAN, J.W. Oligopoly and the theory of games, North Holland, Amsterdam (1977).
- [6] HALMOS, P.R. Measure theory, van Nostrand, Princeton (1959).
- [7] HIMMELBERG, C.J., PARTHASARATHY, T., RAGHAVAN, T.E.S. & VAN VLECK, F.S. Existence of p-equilibrium and optimal stationary strategies in stochastic games, Proc. Amer. Math. Soc. 60, 245-251 (1976).
- [8] KIRMAN, A.P. & SOBEL, M.J. Dynamic oligopoly with inventories, Econometrica 42, 279-287 (1974).



- [9] PARTHASARATHY, T., Discounted, positive and non-cooperative stochastic games, Int. J. of Game Theory 2, 25-37 (1973).
- [10] PARTHASARATHY, T. & STERN, M. Markov games— a survey, University of Illinois, Chicago (1976).
- [11] RIEDER, U. Equilibrium plans for non-zero-sum Markov games, in: Game Theory and Related Topics, D. Moeschlin, D. Pallaschke (eds.), North-Holland, Amsterdam, 91-101 (1979).
- [12] ROGERS, P.D. Nonzero-sum stochastic games, Report ORC 69-8, Operations Res. Center, Univ. of California, Berkeley (1969).
- [13] SHAPLEY, L.S. Stochastic games, Proc. Nat. Acad. Sci. 39, 1095-1100 (1953).
- [14] SOBEL, M.J. Non-cooperative stochastic games, Ann. Math. Stat. 42, 1930-1935 (1971).
- [15] SOBEL, M.J. Continuous stochastic games, J. Appl. Prob. 10, 597-604 (1973).
- [16] TAKAHASHI, M. Equilibrium points of stochastic non-cooperative n-person games, J. Sci. Hiroshima Univ. Ser. A-I 28, 95-99 (1964).
- [17] TIJS, S.H.  $\epsilon$ -Equilibrium point theorems for two-person games, Operations Research Verfahren 26, 755-766 (1977).
- [18] TIJS, S.H. Nash equilibria for non-cooperative n-person games in normal form, Report 7907, Department of Mathematics, Catholic University, Nijmegen, The Netherlands (1979).
- [19] TIJS, S.H. & VRIEZE O.J. Nash equilibria for non-cooperative stochastic games - a survey, Department of Mathematics, Nijmegen, The Netherlands, forthcoming (1979).
- [20] VRIEZE, O.J. The stochastic non-cooperative countable-person game with countable state space and compact action spaces under the discounted payoff criterium, Report BW 66/76, Department of Operations Research, Mathematical Centre, Amsterdam (1976).
- [21] WALD, A. Generalization of a theorem by von Neumann concerning zero sum two person games, Ann. of Math. 46, 281-286 (1945).
- [22] WHITT, W. Representation and approximation of non-cooperative sequential games, Operations Research Center Report, Bell Laboratories, Holmdel, New Jersey (1977).

Address:  
 Department of Mathematics  
 Catholic University  
 Toernooiveld  
 6525 ED Nijmegen  
 The Netherlands